Local Contrast Segmentation to Binarize Images

Marco Block and Raúl Rojas

Freie Universität Berlin

Institut für Informatik

Takustr. 9, 14195 Berlin, Germany

{block,rojas}@inf.fu-berlin.de

Abstract—In this paper, a new binarization algorithm for degraded document images is proposed. The method is based on positive and negative pixel energies using the Laplacian of an image. After a filtering step and morphological operations our local contrast segmentation method is able to detecting connected components. The given approach is applied to cases of sophisticated, challenging documents and other application scenarios like whiteboard and chalk images.

I. Introduction

The early stages of image processing are crucial, since here a lot of information reduction is taking place. A typical example for this aspect is the identification of objects in images [4].

The task of the binarization step is to separate objects as foreground from the rest as background. Therefore, the space of pixel intensities is reduced to $\{0, 1\}$. In this presented case the foreground object is text while the rest is declared as background.

Usage of atomic units, the characters, allows description of the features of text. Characters are in most cases small dark objects, with hard shapes surrounded by a light background.

Thus in small areas of pictures, there may exist large differences of pixel intensities. The sharper the edges, that means a high local contrast, the better the result of detecting text. Usage of this feature is the fundamental idea of the proposed approach for a new binarization algorithm.
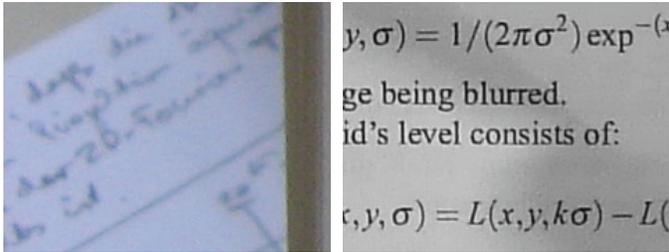

Figure 1. Comparison between sharp and blurry contours. Using blurred contours (left) the OCR delivers no satisfying results. The aim is to interpret this area as background. In contrast, sharp contours (right) should be correctly interpreted as text.

In the left picture of figure 1 there are blurred contours, which should not be recognized as foreground in the binarization step. Only text objects, with high probabilities of being recognized successfully by an OCR shall be identified. In contrast, the example to the right is a typical example, where text can be separated from the background unambiguously.

II. Related Work

The field of binarization algorithms is well-investigated. The objective of a binarization $\mathcal{B}$ is the segmentation of objects of interest in fore- and background. For this aim the result space is set to 1 if a pixel is part of an object (foreground) and 0, if a pixel is not a part of it (background).

There are global methods $\mathcal{B}^G$, that identify a specific threshold Θ for the whole image $\mathbf{I}$ and subsequently set all gray-values of the binarized image $\mathbf{O}$ according to the following presumption

$$\mathbf{O}_{x,y} = \mathcal{B}^G(\mathbf{I}_{x,y}) = \begin{cases} 0 & falls \ \mathbf{I}_{x,y} \leq \Theta \\ 1 & falls \ \mathbf{I}_{x,y} > \Theta \, . \end{cases}$$

There are multiply algorithms to determine this kind of threshold. Well-known methods are those of Abutaleb [23], Kittler [25], Kapur [26], and Otsu [27].

The second class of methods, local binarization $\mathcal{B}^L$, provide for a pixel $\mathbf{I}_{x,y}$ in dependence from a specific neighborhood a local threshold $\Theta_{x,y}$ with

$$\mathbf{O}_{x,y} = \mathcal{B}^L(\mathbf{I}_{x,y}) = \begin{cases} 0 & falls \ \mathbf{I}_{x,y} \leq \Theta_{x,y} \\ 1 & falls \ \mathbf{I}_{x,y} > \Theta_{x,y} \, . \end{cases}$$

This implicates, that local methods compute a threshold function for the whole image. A pixel at the position (x, y) is part of foreground, if the intensity $\mathbf{I}_{x,y} > \Theta_{x,y}$.

Local methods are to be preferred over global methods in most cases, as occurrences of illumination differences prevent the existence of global threshold with ability to separate the fore- and background. Global methods are mainly applied in systems containing an equal illumination, due to their better performance.

A common survey of the different binarization methods can be found in [12].

Subsequently four local binarization methods are described and their results are compared with our new approach in different application scenarios.

A. Method of Niblack

A classical method, introduced by Wayne Niblack in 1986 [24] is often applied and also considered in comparisons [3], [22].

Depending on the mean $\mu(x, y)$ and standard deviation $\sigma(x, y)$ in a small neighborhood around (x, y), the threshold T is computed with

$$T(x, y) = \mu(x, y) + \gamma\sigma(x, y).$$

A neighborhood size of 15×15 pixel and $\gamma = -0.2$ achieves the best results in some experiments [22].

B. Method of Kavallieratou

This method of binarizing document images, introduced by Ergina Kavallieratou in 2005 delivers better results in comparison to other standard methods, i.e. the method of Niblack [10]. The first step is to determine the average illumination T_i of the images computed by

$$T^{(i)} = \frac{\sum_x \sum_y \mathbf{I}_{x,y}^{(i)}}{w(\mathbf{I}) \cdot h(\mathbf{I})},$$

where i is the image of the i-th iteration step of the algorithm. This illumination value will be subtracted from all pixels in step two by

$$\mathbf{S}_{x,y} = 1 - T^{(i)} + \mathbf{I}_{x,y}.$$

In step 3 a histogram equalization is performed on the image $\mathbf{S}$ with

$$\mathbf{I}_{x,y}^{(i)} = 1 - \frac{1 - \mathbf{S}_{x,y}}{1 - E^{(i)}},$$

where $E^{(i)}$ is the minimal pixel value of picture $\mathbf{S}$ during the i-th iteration step.

These 3 steps will be iterated until the document is purged. A document is assumed purged, if no more than 3% of its pixels are transformed. Investigations show that in most cases, 20 is an upper bound for the numbers of iterations [10].

C. Method of Li

The method of Li [9] determines the transition energy of a pixel $\mathbf{I}_{x,y}$ regarding the direct neighbors

$$\nabla^2 \mathbf{I}_{x,y} = \mathbf{I}_{x-1,y} + \mathbf{I}_{x+1,y} + \mathbf{I}_{x,y-1} + \mathbf{I}_{x,y+1} - 4 \cdot \mathbf{I}_{x,y}.$$

This corresponds to a discretized variant of the Laplace filter

$$\nabla^2 \mathbf{I} = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

for an image $\mathbf{I}$ [1]. Subsequently pixels with different prefixes will be identified. These computations take place at the most coarse-grained level of an image pyramid.

The edge points will be localized and removed from the original image, if there are no corresponding edge points in equivalent positions inside the other levels of the pyramid.

Followingly, the pixel pairs which define an edge by reason of changing prefixes, will be summarized in a co-occurence matrix C. The threshold T for the decision of affiliation to fore- and background is given by

$$T = arg\, max_{0 \le t \le L-1} \left(\sum_{m=0}^{t} \sum_{n=t}^{L-1} C_{m,n} \right).$$

D. Method of Ramírez

The main idea of the quantile-linear-algorithm [5] is the combination of an intelligent selection of pixels according to specific properties, analogous to Li, and a statistical examination of local neighborhoods of pixels, analogous to Niblack and Kavallieratou.

The quantile-linear-algorithm is based on computation of the discrete transition energy $\mathbf{E}^D$ of each pixel $\mathbf{I}_{x,y}$ with

$$\mathbf{E}_{x,y}^D(\mathbf{I}) = Max_{\mathbf{I}_{x,y}}^S + Min_{\mathbf{I}_{x,y}}^S - 2\mathbf{I}_{x,y}.$$

Hence, $Min_{\mathbf{I}_{x,y}}^S$ the minimal intensity and $Max_{\mathbf{I}_{x,y}}^S$ the maximal intensity for a small neighborhood S of a pixel $\mathbf{I}_{x,y}$ are calculated. The higher or lower the transition energies of pixels, the more probable it is to relate these pixels to the fore- or background.

The pixels, for which it is not possible to take a decision, are assigned an energy closed to 0. A classification of these pixels will yet have to be determined.

Therefore, first the sets Pos and Neg of positive and negative energies for a positive threshold e with

$$Pos(\mathbf{I},\, e) = \{(x,y)\,|\, \mathbf{E}_{x,y}^D(\mathbf{I}) > e\}$$

and

$$Neg(\mathbf{I},\, e) = \{(x,y)\,|\, \mathbf{E}_{x,y}^D(\mathbf{I}) < e\}$$

for all pixels in the image $\mathbf{I}$ will be defined. Hence, the affiliations of a subset of the image pixels are established. In [5], the image will be subdivided in k rectangular blocks $B_1, B_2, \ldots, B_k$. Inside a block B_i the pixels with high positive energy which satisfy the property

$$|P_i| \le \alpha |\{(x,y)\,|\, \mathbf{E}_{x,y}^D = 0\}|$$

will be identified by means of a parameter α, where $P_i = \{p | p \in Pos(\mathbf{I},\, e) \cap B_i\}$ and $|M|$ is the size of set M. The positive energies of all pixels will be transferred into a histogram. The parameter α decides now, the percentage of pixels with high positive energy that will be used for further computation. For example if $\alpha = 0.1$,10% of the highest energy pixels remain in the set P_i. The negative pixels will be defined analogously.

Mean $\mu_+(P_i)$ and standard deviation $\hat{\sigma}_+(P_i)$ of the gray values inside the block B_i will be computed. In the same way, the mean $\mu_-(N_i)$ and standard deviation $\hat{\sigma}_-(N_i)$ of $N_i = \{n | n \in Neg(\mathbf{I},\, e) \cap B_i\}$ will be computed to define a local threshold with

$$T(B_i) = \frac{\mu_+(B_i) + \mu_-(B_i) + \beta\hat{\sigma}_+(B_i) - \gamma\hat{\sigma}_-(B_i)}{2},$$

where β and γ are tuning parameters.

III. THE NEW APPROACH

Here, a new approach using edge intensities and the properties of the segmented objects without the need of presumptions about expected text sets shall be introduced. Contrary to the quantile-linear-algorithm, which calculates the discrete transition energies, it is possible to efficiently calculate continuous transition energies.

A. Continuous pixel energies

A possibility of calculating continuous local edge intensities of an image $\mathbf{I}$, is to subtract the blurred image $\mathbf{I} * \mathbf{H}^G$, where $\mathbf{H}^G$ is the Gaussian kernel from the original image $\mathbf{I}$ by

$$\mathbf{E}^G(\mathbf{I},\, v) = v \cdot (\mathbf{I} - \mathbf{I} * \mathbf{H}^G).$$

The differences of both images are the continuous edge intensities, where the edge range depends on the used kernel size. With the prefix of scaling parameter v fore- and background can be identified.

The advantage in comparison with the discrete variant of the quantile-linear-algorithm, is beside the accurateness the performance, when using the Fourier transformation [1], [11]. The results can be examined in table I.

image size [MP]	energy discrete [ms]	energy continuous [ms]
1	311	17
2	556	29
4	1301	72
8	5172	281

Table I

The performance test was made on an 3.0 GHz pentium IV processor with 1 GB RAM. Therefore, the discrete and continuous transition energies of 40 different document images of variable sizes and resolutions were computed and the average computation time in milliseconds determined. Determination of the continuous transition energies was 18 times faster.

B. Filtering and adaptation of significant energies

For successful text recognition only significant positive energies $\mathbf{F}^+$ and negative energies $\mathbf{F}^-$ are taken into account for the further computation . The energy thresholds of foreground pixel $\theta_{\varepsilon+}$ and background pixel $\theta_{\varepsilon-}$ will be setted. The positive energies will be determined with $v > 0$ and defined by

$$\mathbf{F}^+_{x,y}(\mathbf{I}) = max(\mathbf{E^G}_{x,y}(\mathbf{I},\, v),\, \theta_{\varepsilon+}).$$

The negative energies will be computed in the same way

$$\mathbf{F}^-_{x,y}(\mathbf{I}) = max(\mathbf{E^G}_{x,y}(\mathbf{I},\, v),\, \theta_{\varepsilon-}),$$

using $v < 0$. The threshold depends on the expected object size, as at this step the blurred image parts are removed, e.g. cases at a relativly grainy resolution. In relation to different application scenarios the parameters $\theta_{\varepsilon+}$ and $\theta_{\varepsilon-}$ have to be set in dependence on the expected object size.

C. Application specific object segmentation

After determining and filtering the translation energies, it is possible to introduce segmenation instructions for different application scenarios. The background images $\mathbf{F}^+(\mathbf{I})$ and $\mathbf{F}^-(\mathbf{I})$ of the image $\mathbf{I}$ generated during the process will be binarized to $\widetilde{\mathbf{F}}^+(\mathbf{I})$ a nd $\widetilde{\mathbf{F}}^-(\mathbf{I})$ with

$$\widetilde{\mathbf{F}}^+(\mathbf{I})_{x,y} = \begin{cases} 1 & ,\, if\ \mathbf{F}^+_{x,y} > 0 \\ 0 & ,\, else \end{cases}$$

and

$$\widetilde{\mathbf{F}}^-(\mathbf{I})_{x,y} = \begin{cases} 1 & ,\, if\ \mathbf{F}^-_{x,y} > 0 \\ 0 & ,\, else \end{cases}.$$

Afterwards, the background will be expended using the morphological operation dilation, to close small gaps. Those pixels of the foreground which were previously set to 1 in $\widetilde{\mathbf{F}}^+$ will now be considered background and set to 0 by

$$\mathbf{R}(\mathbf{I})_{x,y} = \begin{cases} 0 & ,\, if\ \widetilde{\mathbf{F}}^+(\mathbf{I})_{x,y} = 1 \\ \mathcal{M}^D(\widetilde{\mathbf{F}}^-(\mathbf{I}))_{x,y} & ,\, else \end{cases}.$$

The requested objects of interest are embedded into connected background areas in $\mathbf{R}(\mathbf{I})$ and can now be determined with standard methods to identify connected components [19].
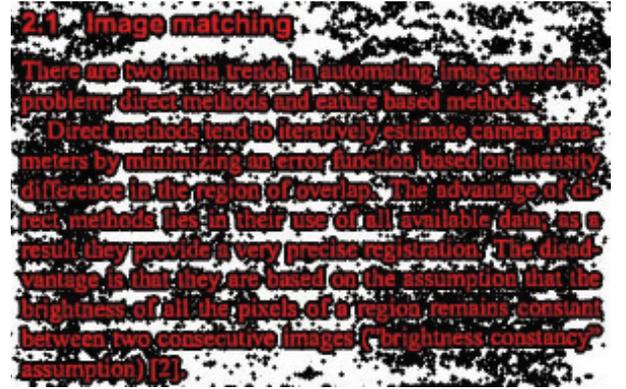

Figure 2. The objects of interest (red) in the image are completly embedded into background pixels ($\mathbf{R}(\mathbf{I})$). The connected components can be extracted now.

Undesired edges, for example strong shadows, will not be considered in this method, because the objects of interests are not completely enbedded into background pixels. The reason is that the edges in the images are totally or partially blurred.

Connected components $C_{\mathbf{R}(\mathbf{I})}$ of the image $\mathbf{R}(\mathbf{I})$ that were determined as foreground objects relate to the object candidates with

$$C_{\mathbf{I}} = \{(x,y) | 0 \leq x < w(\mathbf{I}),\, 0 \leq y < h(\mathbf{I})\},$$

where $w(\mathbf{I})$ is the width and $h(\mathbf{I})$ the height of an image $\mathbf{I}$. With the sizes of the requested objects it is possible to draw a conclusion to the set of connected components $C_{\mathbf{R}(\mathbf{I})}$ with

$$C^X_{\mathbf{R}(\mathbf{I})} = \{c | c \in C_{\mathbf{R}(\mathbf{I})}, \theta^{min}_X \leq |c| \wedge |c| \leq \theta^{max}_X\},$$

where θ^{min}_X is the minimal and θ^{max}_X the maximal size of the object of interest X. The object size is determined by the number of pixels.

The pixels of the identified connected components $C^X_{\mathbf{R}(\mathbf{I})}$ define the foreground areas and will be set to 1. The other pixels will be set to 0. For the requested text it is reasonable to set the foreground to black and the background to white. A comprehensive example is shown in figure 3. This methodology will be labeled in the following as *Local Contrast Segmentation* (LCS).
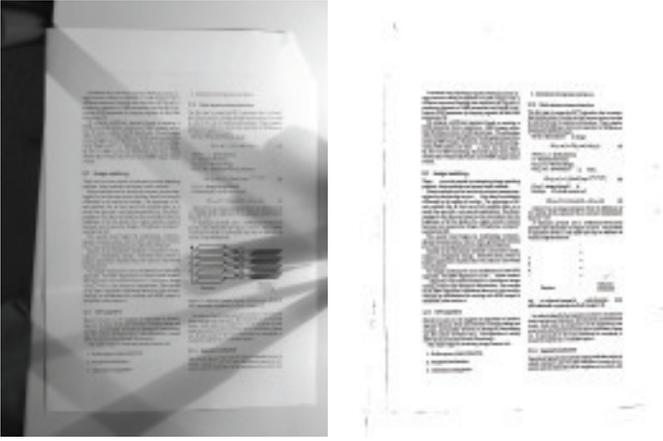
Figure 3. A document with a lot of shadows was used (left) and binarized with the introduced method (right). It was even possible to remove the contained figure.

IV. EXPERIMENTS AND RESULTS

In the following sections, three different application scenarios for binarization are presented: text on a whiteboard, text on a chalkboard and documents taken by a handheld camera.

A. Scenario: text on a whiteboard

There are already some publications of text recognition concerning whiteboards [6], [14], [18]. The work of Markus Wienecke firstly uses average intensity values of the textfree whiteboard and secondly collects information from different images [14]. The electronic pen recording system eBeam interface [30] was used in the work of Markus Liwicki [6].

The segmentation of written text on a whiteboard is an easy task for the LCS-algorithm, because the whiteboard crayon is producing a strong contrast on the background.

Images with a resolution of 6 megapixel were taken with a Canon Powershot A700 digital camera. Small and large areas of the image can be of interest, therefore the parameters will be set to $\theta_{Whiteboard}^{min} = 1$ and $\theta_{Whiteboard}^{max} = \infty$. The best experimental results were achieved with $\theta_{\varepsilon+} = 30$ and $\theta_{\varepsilon-} = 1$.

In this way the text can be segmented without loss of information. Figure 4 shows an example image and the related output of the LCS-algorithm.
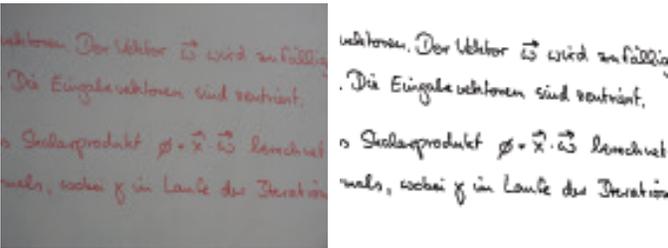

Figure 4. Output of LCS-algorithm (right) of an whiteboard image (left)

1) Experimental results: Inside the handwritten texts there are some images which exact binarization results are not easily automatic validated, therefore human validation is used analogous to other fields of image binarization [7].

Test persons were consulted for the evaluation. Their task was to choose the best result from 20 different examples of binarizations. As quality criteria the complete result without redundant parts was used. Figure 5 shows the resulting diagram.
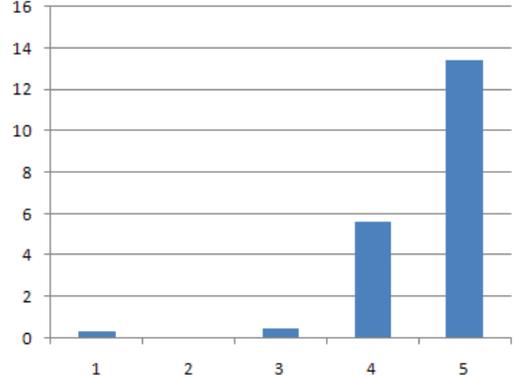

Figure 5. Methods: Niblack (1), Kavallieratou (2), Li (3), quantile-linear (4) and LCS (5). For the evaluation of handwritten text on whiteboards the segmentation method was found to be best performing binarization method.

In 13 of 20 cases the 23 test persons determined the segmentation method as the best performing choice. An exception was given for images with small contrast between the crayon color (e.g. green) and the board or when an old pen was used. In these cases, the quantile-linear-algorithm was preferred. In rare cases Niblack and Kavallieratou delivered better results.

The test images are available online [32].

B. Scenario: text on a chalkboard

So far it was assumed that the foreground is darker related to background like in most of text images. In some scenarios this is not the case. The chalkboard works with text on a darker background. Accordingly, the segmentation algorithm should be extended marginally.

Instead of only segmenting the foreground, the foreground image will be extended with dilation and covered with the background

$$\widetilde{\mathbf{R}}(\mathbf{I})_{x,y} = \begin{cases} 0 & , falls\ \widetilde{\mathbf{F}}^-(\mathbf{I})_{x,y} = 1 \\ \mathcal{M}^D(\widetilde{\mathbf{F}}^+(\mathbf{I}))_{x,y} & , sonst \end{cases}.$$

Now the connected components in $\widetilde{\mathbf{R}}(\mathbf{I})$ will be identified and defined as foreground, the rest as background. The application of the segmentation method achieves better results in comparison to the alternatives on clean contours (figure 6).
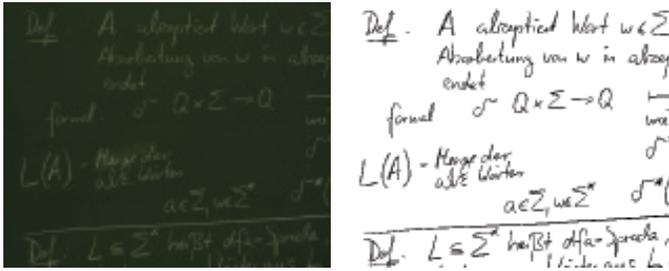
Figure 6. Application of the segmentation method on a chalkboard image

Figure 8. In the left image the illumination is bad, some parts of the text are blurred lacking contrast. The binarization only identifies objects with sharp contours (right).

Chalk text is identified analogously to text on a whiteboard, therefore the parameters $\theta_{Kreide}^{min} = 1$ and $\theta_{Kreide}^{max} = \infty$ were chosen. As a result there is no restriction of the resulting connected components from the image $\widetilde{\mathbf{R}}(\mathbf{I})$.

In experiments, the parameters $\theta_{\varepsilon+} = 20$ and $\theta_{\varepsilon-} = 1$ delivered the best results. For all the other binarization algorithms the image was primarily inverted.

1) Experimental results: To predict the exact positions of the text on a board or to construct a benchmark for it, is nearly impossible. Therefore, evaluations by the test persons were used. The evaluation results for 20 examples by 31 test persons are shown in figure 7.
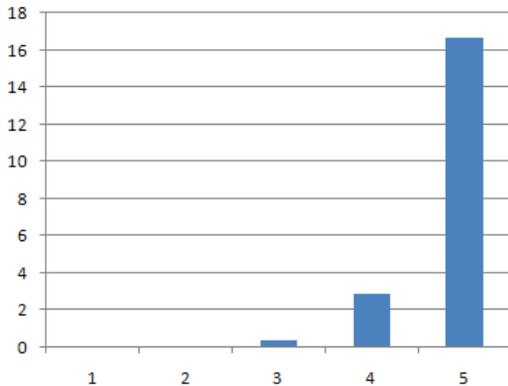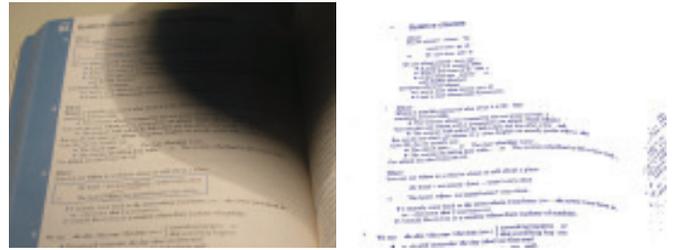

Figure 7. Methods: Niblack (1), Kavallieratou (2), Li (3), quantile-linear (4) and LCS (5). Again the segmentation method provides the best results.

In this application scenario, the segmentation method displays its strength and outperforms the other methods. The quantile-linear-algorithm completes the identified parts of text in a better way, improving their readability.

The figures of the test and the evaluation are available online [33].

C. Scenario: text documents

Analogous the segmentation of text from the whiteboard gaining excellent binarization results from images of text documents using segmentation method. Blurred parts will be removed (figure 8).

If the text document focussed by the camera, then the local contrast will be located by the segmentation method. Most of the background objects will be eliminated based on their defocussation. This approach is a good alternative for applications based on detection of text documents [13], [31].

The setting of parameters $\theta_{Textdokument}^{min}$, $\theta_{Textdokument}^{max}$, $\theta_{\varepsilon+}$ and $\theta_{\varepsilon-}$ depends on the used camera and the expected text size related to the distance.

Working with a 2 megapixel webcamera QuickPro 9000 by Logitech with properties of autofocus and Carl-Zeiss optics [29] opening the possibility to successfully recognize text with font sizes of 10 to 28 inside a distance range of 10 to 40 cm.

The interpolated 8 megapixels by the QuickPro 9000 software are not sufficient to furtherly extend the distance. The parameters $\theta_{Textdokument}^{min} = 2$, $\theta_{Textdokument}^{max} = 5000$, $\theta_{\varepsilon+} = 10$ and $\theta_{\varepsilon-} = 1$ were experimentally determined to be the best in practice.

1) Experimental results: The verification of segmentation of known text parts is easier to perform than that of the previous applications.

Hence, we prepared a database with different images of 40 well-known text parts variating in font type and size. The images were binarized by the five methods of Niblack, Kavallieratou, Li, quantile-linear and LCS and the resulting images were processed by the OCR Tesseract [28] for text detection.

The results of the experiment were compared to the five recognized texts using the Needleman-Wunsch-algorithm. The recognition rates are shown in figure 9.
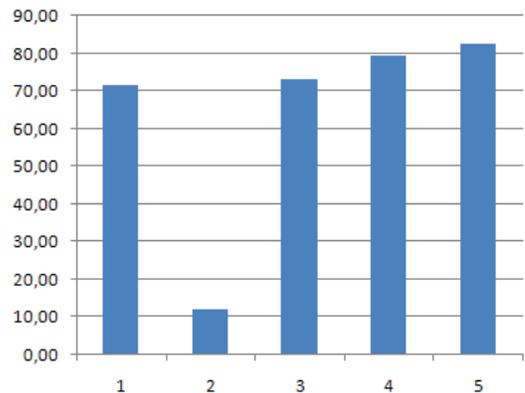

Figure 9. Used methods: Niblack (1), Kavallieratou (2), Li (3), quantile-linear (4) and LCS (5). The LCS method performed best in the experiment.

The 40 used text parts in the experiment were written in German. However unexpectedly, Tesseract had a lot of difficulties concerning identification of quotation marks. In these cases the OCR regularly detected two commas. The single way to solve this problem was to completely filter out any punctuation marks.

In comparison to the other methods LCS delivered better results.

V. Conclusion and Future Work

The application of the new binarization method is specialized on hand-written and printed text. The presented binarization method shall be compared to more methods in additional experiments. It is also interesting to take into consideration commercial OCRs other than Tesseract.

References

[1] R.C. Gonzales, R.E. Woods, „*Digitial Image Processing*", 3.Auflage, ISBN 978-0131687288, Pearson Verlag 2008
[2] A.K. Jain, P. Flynn, A.A. Ross, „*Handbook of Biometrics*", ISBN 978-0387710402, Springer Verlag 2008
[3] Y. Xi, Y. Chen, Q. Liao, „*A Novel Binarization System for Degraded Document Images*", 9th International Conference on Document Analysis and Recognition (ICDAR 2007), IEEE Computer Society 2007, ISBN 978-0-7695-2822-9, Vol.1, pp.287-291, Curitiba/Brazil, 2007
[4] J. Shotton, A. Blake, R. Cipolla, „*Multi-Scale Categorical Object Recognition Using Contour Fragments*", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI 2007), Vol.30, No.7, pp. 1270-1281, 2008
[5] M. Ramírez, E. Tapia, M. Block, R. Rojas, „*Quantile Linear Algorithm for Robust Binarization of Digitilized Letters*", 9th International Conference on Document Analysis and Recognition (ICDAR 2007), IEEE Computer Society 2007, ISBN 978-0-7695-2822-9, Vol.2, pp.1158-1162, Curitiba/Brazil, 2007
[6] M. Liwicki, H. Bunke, „*Handwriting Recognition of Whiteboard Notes - Studying the Influence of Training Set Size and Type*", International Journal of Pattern Recognition and Artificial Intelligence, Vol. 21, No. 1, pp. 83-98, 2007
[7] X. Liang, „*Fingerprint Image Analysis Using Computational Geometric Techniques*", Dissertation in School of Information Science, Japan 2006
[8] B. Gatos, I. Pratikakis, S.J. Perantonis, „*Adaptive degraded document image binarization*", Pattern Recognition 39, pages 317-327, 2006
[9] Y. Li, C. Suen, M. Cheriet, „*A Threshlod Selection Method Based on Multiscale and Graylevel Co-occurrence Matrix Analysis*", Eighth International Conference on Document Analysis and Recognition (ICDAR 2005), pp. 463-467, Seoul/Korea, 2005
[10] E. Kavallieratou, „*A Binarization Algorithm specialized on Document Images and Photos*", Eighth International Conference on Document Analysis and Recognition (ICDAR 2005), pp. 463 - 467, Seoul/Korea, 2005
[11] K.D. Tönnies, „*Grundlagen der Bildverarbeitung*", ISBN 978-3827371553, Pearson Studium 2005
[12] M. Sezgin, B. Sankur, "*Survey over image thresholding techniques and quantitative performance evaluation*", Journal of Electronic Imaging 13(1), pp.146–165, January 2004
[13] E. Bertucci, M. Pilu, M. Mirmehdi, „*Text Selection by Structured Light Marking for Hand-held Cameras*", IEEE/IAPR 7th International Conference on Document Analysis and Recognition, ISBN 0-7695-1960-1, pp. 555–559, 2003
[14] M. Wienecke, G.A. Fink, G. Sagerer, „*Towards Automatic Video-based Whiteboard Reading*", Int. Journal on Document Analysis and Recognition, Vol.7, pp.87-91, 2003
[15] D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar, "*Handbook of Fingerprint Recognition*", ISBN 978-0-387-95431-8, Springer-Verlag 2003
[16] I.K. Kim, D.W. Jung, R.H. Park, „*Document image binarization based on topographic analysis using water flow model*", Pattern Recognition 35, pp. 265-277, 2002
[17] N. Papamarkos, „*A Technique for Fuzzy Document Binarization*", Proceedings of the 2001 ACM Symposium on Document engineering, pp.152-156, 2001
[18] E. Saund, „*Method for segmenting handwritten lines of text in a scanned image*", European Patent: EP19950306118, 2001
[19] Stockman G., Shapiro L.G.: „*Computer Vision*", ISBN 978-0130307965, Addison Wesley Verlag, 2001
[20] J. Sauvola, M. Pietikainen, „*Adaptive document image binarization*", Pattern Recognition, Vol.33, Issue 2, pp. 225-236, 2000
[21] D. Maio, D. Maltoni, „*Direct Gray-Scale Minutiae Detection in Fingerprints*", IEEE Transactions on Pattern Analysis Machine Intelligence, Vol.19, No.1, pp.27-40, 1997
[22] O.D. Trier, A.K. Jain, „*Goal-Directed Evaluation of Binarization Methods*", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol 17, pp.1191-1201, 1995
[23] A.S. Abutaleb, „*Automatic thresholding of gray-level pictures using two-dimensional entropy*", Computer Vision, Graphics and Image Processing, Vol.47, pp.22-32, 1989
[24] W. Niblack, „*An Introduction to Digital Image Processing*", Prentice Hall, ISBN 978-0134806747, Englewood Cliffs, NJ, pp.115-116, 1986
[25] J. Kittler, J. Illingworth, „*Minimum error thresholding*", Pattern Recognition, Vol.19, No.1, pp.41-47, 1986
[26] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, „*A new method for gray-level picture thresholding using the entropy of the histogram*", Computer Vision, Graphics and Image Processing, Vol.29, pp.273-285, 1985
[27] N. Otsu, „*A threshold selection method from gray-level histograms*", IEEE Trans. Systems Man Cybernet 9 (1), pp.62-66, 1979
[28] Tesseract: *http://code.google.com/p/tesseract-ocr/*
[29] Webseite von Logitech: *http://www.logitech.com*
[30] Webseite des E-Beam-Systems: *http://www.e-beam.com/*
[31] Indoor-Scene-Projektseite: *http://www.cs.bris.ac.uk/Research/Digitalmedia/docum.html*
[32] WhiteBoard-Evaluation: *http://page.mi.fu-berlin.de/block/WhiteBoard.pdf*
[33] Kreidetafel-Evaluation: *http://page.mi.fu-berlin.de/block/KreideTafel.pdf*